



## Commentary

**Cite this article:** Pouw W, Dixon JA. 2022  
What you hear and see specifies the perception  
of a limb-respiratory-vocal act. *Proc. R. Soc. B*  
**289:** 20221026.  
<https://doi.org/10.1098/rspb.2022.1026>

Received: 26 May 2022

Accepted: 27 June 2022

**Subject Category:**

Neuroscience and cognition

**Subject Areas:**

biomechanics, behaviour, evolution

**Author for correspondence:**

Wim Pouw

e-mail: [wim.pouw@donders.ru.nl](mailto:wim.pouw@donders.ru.nl)

# What you hear and see specifies the perception of a limb-respiratory-vocal act

Wim Pouw<sup>1,2</sup> and James A. Dixon<sup>3</sup>

<sup>1</sup>Donders Institute for Brain Cognition and Behaviour, Radboud University, Nijmegen, Gelderland, The Netherlands

<sup>2</sup>Max Planck Institute for Psycholinguistics, Nijmegen, Gelderland, The Netherlands

<sup>3</sup>Department of Psychology, University of Connecticut, Storrs, CT, USA

A commentary on 'Beat gestures influence which speech sounds you hear' [1].

Bosker & Peeters [1] report an extensive and well-executed demonstration of how perception of vocalic aspects of speech is a multimodal affair that can be instantiated by visual information about co-speech upper limb movement. This discovery of a 'gesture McGurk effect' (see also [2]) can be considered as a fascinating extension of the classic McGurk effect which is originally obtained in relation to articulatory aspects of speech perception. Instead, the authors interpret their discovery as a logically distinct phenomenon.

Bosker and Peeters seem to hold that gesture and speech are causally *independent* modes of communication as the gesture McGurk effect does not reflect information originating from 'the same communicative channel (i.e. articulation)' [1, p. 17] as is the case for the classic McGurk effect. The authors follow this through by accepting that the *dependence* of speech perception with gesture is achieved via 'top-down' cognitive inference: 'As such, what we perceive is the model of reality that our brains provide us by binding visual and auditory communicative input, and not reality itself' [1, p. 7].

To address the first issue, there is evidence that respiratory-vocalic aspects of speech such as the fundamental frequency and intensity are directly modulated by physical impulses that are produced by beat-like upper limb movements [3–6]. This modulation is attributed to upper limb movements recruiting a wider ensemble of posture-maintaining muscles around the trunk which are implicated with control of expiration [7–9] and thus vocalic aspects of speech [5,10]. What this means is that there is a causally dependent biophysical relationship between vocalic aspects of speech and upper limb movements. For the present gesture McGurk effect this means that beat gestures and vocalic aspects of speech have a causal connection in reality. Specifically, markers of lexical stress that are affected by rate of expiratory flow, such as intensity and F0, will be affected because of gesture-induced changes in sub-glottal pressure (for a detailed account see [11]). Given that the rate of expiratory flow is less directly related to duration we also think it is less obviously connected to the biomechanical impulses that gestures generate.

The experimental manipulation in [1] is such that the real sound source and the gesture source are detached so as to be manipulated. Yet, this does not mean that perception of vocalic aspects of speech as affected by seeing gesture is therefore detached from reality; the perception is rooted in the real connection of gesture and speech. Echoing classic critiques of cognitivist interpretations of the classic McGurk effect, listeners are not attending to the sound alone they are perceiving a limb-vocalic speech act, and varying information about physical impulses of gesture interacts with audition in the perception of a more global array of multimodal information [12,13].

To conclude, this commentary is meant to guard against an overly cognitively complex interpretation of the relation between gesture and speech, both in perception and production. We have argued that this discovery of the gesture McGurk effect beautifully reflects that vocalic actions tend to be perceived by attuning to 'simultaneous changes in the structure of multiple

forms of ambient energy' [12 p. 196] much like the articulatory McGurk effect.

**Data accessibility.** This article has no additional data.

**Authors' contributions.** W.P.: conceptualization, writing—original draft, writing—review and editing; J.A.D.: writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare we have no competing interests.

**Funding.** This research has been funded by a VENI grant (VI.Veni.201G.047) awarded by the Dutch Research Council (NWO) to Wim Pouw (PI).

## References

1. Bosker HR, Peeters D. 2020 Beat gestures influence which speech sounds you hear. *Proc. R. Soc. B* **288**, 20202419. (doi:10.1098/rspb.2020.2419)
2. Bujok R, Meyer A, Bosker HR. 2022 Audiovisual perception of lexical stress: beat gestures are stronger visual cues for lexical stress than visible articulatory cues on the face. *PsyArXiv* (doi:10.31234/osf.io/y9jck)
3. Pouw W, de Jonge-Hoekstra L, Harrison SJ, Paxton A, Dixon JA. 2020 Gesture-speech physics in fluent speech and rhythmic upper limb movements. *Ann. NY Acad. Sci.* **1491**, 89–105. (doi:10.1111/nyas.14532)
4. Pouw W, Harrison SJ, Dixon JA. 2019 Gesture-speech physics: the biomechanical basis of the emergence of gesture-speech synchrony. *J. Exp. Psychol. Gen.* **149**, 391–404. (doi:10.1037/xge0000646)
5. Pouw W, Harrison SJ, Esteve-Gibert N, Dixon JA. 2019 Energy flows in gesture-speech physics: exploratory findings and pre-registration of confirmatory analysis. *PsyArXiv* (doi:10.31234/osf.io/c7456)
6. Pouw W, Paxton A, Harrison SJ, Dixon JA. 2020 Acoustic information about upper limb movement in voicing. *Proc. Natl Acad. Sci. USA* **117**, 11 364–11 367. (doi:10.1073/pnas.2004163117)
7. Cordo PJ, Nashner LM. 1982 Properties of postural adjustments associated with rapid arm movements. *J. Neurophysiol.* **47**, 287–302. (doi:10.1152/jn.1982.47.2.287)
8. Hodges PW, Sapsford R, Pengel LHM. 2007 Postural and respiratory functions of the pelvic floor muscles. *Neurorol. Urodyn.* **26**, 362–371. (doi:10.1002/nau.20232)
9. Hodges PW, Gandevia SC. 2000 Activation of the human diaphragm during a repetitive postural task. *J. Physiol.* **522**, 165–175. (doi:10.1111/j.1469-7793.2000.t01-1-00165.xm)
10. Finnegan EM, Luschei ES, Hoffman HT. 2000 Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *J. Speech Lang. Hear. Res.* **43**, 934–950. (doi:10.1044/jslhr.4304.934)
11. Pouw W, Fuchs S. 2021 Origins of vocal-entangled gesture. *PsyArXiv* (doi:10.31234/osf.io/egnar)
12. Stoffregen TA, Bardy BG. 2001 On specification and the senses. *Behav. Brain Sci.* **24**, 195–213. (doi:10.1017/S0140525X01003946)
13. Fowler CA, Dekle DJ. 1991 Listening with eye and hand: cross-modal contributions to speech perception. *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 816–828. (doi:10.1037/0096-1523.17.3.816)